Department of Business Administration

UZH Business Working Paper Series

Working Paper No. 346

The Impact of High School Exit Exams on Graduation Rates and Achievement

Katherine Caves and Simone Balestra

September 2014

University of Zurich, Plattenstrasse 14, CH-8053 Zurich,
http://www.business.uzh.ch/forschung/wps.html

UZH Business Working Paper Series

Contact Details

Katherine Caves

University of Zurich

Department of Business Administration

Plattenstrasse 14, CH-8032 Zurich, Switzerland

katherine.caves@business.uzh.ch

Tel.: +41 44 634 29 29

Simone Balestra

University of Zurich

Department of Business Administration

Plattenstrasse 14, CH-8032 Zurich, Switzerland

simone.balestra@business.uzh.ch

Tel.: +41 44 634 42 20

The Impact of High School Exit Exams on Graduation Rates and Achievement*

Katherine Caves† and Simone Balestra‡

September 2014

PRELIMINARY — DO NOT CITE WITHOUT PERMISSION

Abstract

In this paper, we examine the long-term effects of high school exit exams (HSEEs) on graduation rates and achievement using an interrupted time series approach. We find that introducing a HSEE has an overall positive effect on graduation rate trends, an effect which is heterogeneous over time. In the year of introduction and the following three years we find a negative impact of HSEE on graduation rates; this negative impact is short-lived and becomes positive over the long term. We perform robustness checks using states that do not have HSEEs as control group. We also estimate a pre-intervention negative effect, suggesting that high schools start preparing for the HSEE before its actual introduction. We find no effects for achievement, possibly due to the lack of meaningful cross-state achievement data in the time period studied.

Keywords: High school exit exams, graduation rates, achievement.

JEL Classification: C33, I21.

---

* We are grateful to Uschi Backes-Gellner, Simon Janssen, Matthias Krapf, Christian Rupietta, and Andrea Schenker-Wicki for their constructive comments.

† Research associate at the Graduate School of Business at the University of Zurich, Switzerland. Corresponding author: katherine.caves@business.uzh.ch

‡ Research associate at the Graduate School of Business at the University of Zurich, Switzerland.

1 Introduction

As part of the increasing trend towards school accountability and standards-based education over the past two decades, most American states have implemented high school exit exam (HSEE) policies requiring that all high school students pass a test to graduate (Reardon, Arshan, Atteberry, and Kurlaender, 2010). Exit exams ideally ensure a minimum achievement level for high school graduates and raise the value of high school diplomas, but many scholars worry about the negative effects of such exams on student motivation, graduation rates, and equity (Dee and Jacob, 2006). Estimation of the real effects of high school exit exams is difficult because randomized experimental application of the policy does not exist, so disagreement remains on whether exit exams have positive effects—increasing the achievement level and degree value of high school graduates—or negative effects—decreasing graduation rates and fostering inequity. Although discussion of exit exams' effects often centers on graduation rates, it remains unclear whether or not requiring an exam negatively affects the graduation prospects of high school students.

In this paper, we exploit the staggered implementation of HSEEs at the state level over the period of 1990-2013 to examine their real long-term effects on graduation rates and achievement and how those effects persist over time. We center each state's time line on the year in which high school diplomas were first withheld based on exam scores. In the linear specification, we find a downward jump in graduation rate in the first year of withholding that recovers completely within five years. In the non-parametric specification, we find a short-term decrease in graduation rates both preceding the application of the exam and immediately following, but graduation losses are recovered within four years and continue improving due to the slope of the graduation rate trend increasing after the HSEE.

Previous findings indicate that the implementation of exam policies increases dropout rates, especially for minority and low-income students (Bishop and Mane, 2001; Dee and

Jacob, 2006; Jacob, 2001), although Warren and Edwards (2005) find no effects. The average effects of educational policies including HSEEs can be misleading as they often have distributional effects (Jackson and Page, 2013; Ou, 2010). Dee and Jacob (2006) find that exams reduce the probability of completion overall, but that these effects are particularly strong for Black students and in urban school districts with high levels of poverty or minority enrollments. At the same time, they find that HSEEs actually lower dropout rates in more affluent districts.

For graduation as opposed to dropout rates, the finding of no effects is more prevalent (Carnoy and Loeb, 2002; Greene and Winters, 2005; Grodsky, Warren, and Kalogrides, 2009; Warren and Edwards, 2005), again with some exceptions (Amrein and Berliner, 2002; Marchant and Paulson, 2005). Establishing causality, however, is difficult and there have been some methodological shortcomings; work focused on causal analysis finds a slight increase in dropout rates among black and low-income students with HSEEs (Dee and Jacob, 2006; Murnane, 2013; Warren, Jenkins, and Kulick, 2006). Overall, there is some evidence that HSEEs can improve student achievement and the attainment of high school diplomas, but this effect is highly dependent on school resources and subject to equity issues.

We begin by introducing the theory surrounding exit exams (Section 2) and previous findings (Section 3). In Section 4 we describe our data and the staggered implementation of exit exams in the United States. We demonstrate our methodological approach in Section 5 and report results in Section 6. We conclude and discuss potential mechanisms for these effects in Section 7.

2 Theoretical Background

The effects of HSEEs on graduation rates can be approached from two distinct theoretical perspectives: economics of education and sociology of education. From an economic theory perspective, increasing the difficulty of acquiring any qualification will decrease the number of individuals achieving that qualification. Importantly, qualifications gain value as a signal of ability when they are difficult to attain (Arcidiacono, Bayer, and Hizmo, 2010; Tyler, Murnane, and Willet, 2000; Spence, 1973). From a sociological and pedagogical point of view, the effects of HSEEs are less clear. Exams could potentially enhance student achievement by focusing school curriculum and teacher instruction on the relevant standards, encouraging the provision of targeted assistance for low achievers, and motivating students (Bishop, 1997; Bishop, Mane, and Bishop, 2001). Conversely, the exam could demotivate lower-achieving students, incentivize schools and teachers to give up on “hopeless cases,” and enhance inequity especially among non-English-speaking students (Booher-Jennings, 2005). If the HSEE is based on flawed standards or fails to adequately measure standards, all of these issues are compounded.

Economics predicts positive overall effects of HSEEs for the value of a high school degree and the future prospects of graduates, but acknowledges that these benefits come at a cost of lower attainment. The predicted benefits of HSEEs are non-negligible for both graduates and as a tool for school policy. Achievement of a known minimum standard creates accountability and meaning for a high school education. The increased standards and achievement raise the value of the diploma as a signal of graduates’ ability, which improves the labor market and college application prospects of graduates by reducing informational asymmetries in the labor and college market. At the school policy level, the HSEE provides a means of measuring school performance, making schools accountable for teaching the material for which they are responsible (Bishop and Mane, 2001; Reardon, Arshan, Atteberry, and Kurlaender, 2010). The state can be sure that tested schools will work towards achieving the standards on which they are tested.

However, these benefits are inextricable from lower attainment: an HSEE that prevents the lowest achievers from graduating will increase average achievement among high school graduates by mathematical necessity even if it has no effects on the behavior of students or schools. Increased standards raise performance, but at least part of this effect comes from increased selection. HSEEs are an advisable policy from an economics of education standpoint, but they do prevent the lowest achievers from graduating.

The sociological perspective on HSEEs is more behavioral than incentivist in its predictions for how schools, teachers, and students will react to the implementation of an HSEE and is more skeptical of the accuracy of measurement and appropriateness of standards contained in the exam. As a result, the value of an HSEE and its likely effects on graduation rates are more complex and unpredictable. Positive outcomes of HSEEs revolve around their ability to focus school, teacher, and student efforts (Bishop and Mane, 2001; Bishop, Mane, and Bishop, 2001). Schools and teachers that know students will be tested on specific material will orient curriculum and instruction around ensuring that students master that content. Students at risk of failing can be given targeted assistance, which can include language services for students whose native language is not English. Students, aware that they will be held responsible for what they learn in class, may be more motivated to study and retain key concepts.

Alternatively, however, HSEEs can generate negative incentives that undermine school, teacher, and student behavior. Schools and teachers may take content specificity too far, yielding curriculum and instruction that covers only what is on the test without substantive context or meaning. Teachers may decide that some students have no hope of passing and prioritize them lower than those on the bubble (Booher-Jennings, 2005). Students who believe they cannot pass may be demotivated or incentivized to drop out. Regardless of the behaviors of schools, teachers, and students, lack of resources in some schools and districts might prevent the implementation of necessary changes, creating a social justice problem (Plunk, Tate, Bierut, and Grucza, 2014). There are strong theoretical arguments for both positive and negative effects of HSEEs on school, teacher, and student behavior and performance.

Standards or exams that do not adequately represent or measure appropriate educational goals are an enormous obstacle to the success of HSEEs from any perspective. Educational standards are difficult to set, especially given that the purpose of education is contested in the United States. A successful high school education can be defined as preparing student for any or all of entering the labor market, attending college, acting as an informed citizen, and functioning in society. States define curricula for the material and level of mastery required of their high school graduates, and there is no guarantee that these standards or the goals they represent are necessary or sufficient for graduates to succeed in later life. This is further complicated by the potential for HSEEs to fail to accurately measure achievement of standards. Schools, teachers, and students are incentivized to pass the exam, but the skills necessary for that goal may not match the curriculum standards or relevant skills if the HSEE itself is poorly conceived (Akerlof, 1970; Gibbons and Katz, 1991).

HSEEs focus school, teacher, and student energy on mastery of whatever skills and material are required to pass. Exams can incentivize schools and teachers to refine curriculum and instruction and to target students who need the most help. Conversely, they might encourage narrow rote learning and the neglect of students perceived to have no chance at passing. Students might be motivated to study or discouraged from attempting a test they believe they cannot pass. All of this can mean a higher-quality graduating class with more valuable diplomas, but it may come at a cost of lower graduation rates—especially among students who do not speak English as their first language or who teachers may be more likely to discount. While exit exams will always need refinement to ensure that they adequately measure appropriate standards, the first concern with the tests at a policy level is equity, for which the first indication would be a significant drop in overall graduation rates.

3 Previous Findings on Exit Exams

Prior studies of HSEEs have examined their effects on dropout, completion, and graduation rates, with some additional work on their impacts on achievement and student educational trajectories.

More recent studies of HSEEs and dropout emphasize heterogeneous effects across student groups and the potential for exams to increase inequity. HSEEs tend to increase dropout rates, especially in disadvantaged populations (Bishop and Mane, 2001; Dee and Jacob, 2006; Jacob, 2001). The average effects of educational policies can be misleading as they often have distributional effects (Bitler, Gelbach, and Hoynes, 2006). This is highlighted by Ou (2010), who uses a regression discontinuity design around the margin of barely failing and passing the New Jersey HSEE to find that while barely failing students are generally more likely to drop out than barely passing students, the negative effect is strongest for minority and low-income students. Dee and Jacob (2006) find that HSEEs reduce the probability of completion overall, but that these effects come from black students and school districts with high levels of poverty or minority enrollments in urban areas, while the HSEE actually lowers dropout rates in more affluent districts. HSEEs have the potential to enhance attainment by focusing instruction and student effort, but only in the presence of adequate resources. More critically, these potential positive effects are apparently weaker than the systematic negative effects of poverty and minority status in American education (Plunk, Tate, Bierut, and Grucza, 2014).

HSEEs themselves can shape the educational trajectories of students. Evidence from a study on Turkish data indicates that high-stakes examinations may actually help reduce achievement gaps based on student background by promoting learning over the course of multiple re-takings (Frisancho, Krishna, Lychagin, and Yavas, 2013). This supports the intuition that HSEEs may motivate students and help focus their efforts on mastery of relevant material, but also highlights the importance of designing an exam that measures relevant content. Reardon, Arshan, Atteberry, and Kurlaender (2010) examine the impact of failing an HSEE in 10$^{\text{th}}$ grade—with two years remaining to pass before graduation—and find that barely failing the exam has no effect on students' academic trajectories, course taking, or graduation probability except for the very lowest achievers. Those authors conclude that negative effects of HSEEs on graduation rates come exclusively from the very lowest achievers. This supports the assertion that HSEEs may have overall positive effects, preventing from graduation only those students who may not be prepared to earn a diploma at all. Still, this does not address the equity issues that have been empirically demonstrated in the racial and socioeconomic distribution of these effects.

When the focus is turned to graduation rather than dropout rates, the finding of no effects is more prevalent (Carnoy and Loeb, 2002; Greene and Winters, 2005; Grodsky, Warren, and Kalogrides, 2009; Warren and Edwards, 2005), again with some exceptions (Amrein and Berliner, 2002; Marchant and Paulson, 2005). The isolation of HSEEs as the cause of these reported effects, however, is difficult and there have been some methodological shortcomings; work intended to remedy those issues finds a slight increase in dropout rates among black and low-income students when HSEEs are in place (Dee and Jacob, 2006; Murnane, 2013; Warren, Jenkins, and Kulick, 2006). Overall, there is some evidence that HSEEs can improve student achievement and the attainment of high school diplomas, but that this is highly dependent on school resources and subject to major racial and socioeconomic equity issues.

Graduation rates in general have been steadily improving throughout the first decade of the 21st century, especially among certain student groups. In a review of the topic, Murnane (2013) outlines the rise of graduation rates through the 1900s until their stagnation in the last three decades of that century, then their recent improvement. The recent rise in graduation rates appears to come largely from major increases in high school graduation among black and Hispanic students, an increase that has occurred simultaneously with the implementation of HSEE policies by a number of states. It remains unclear why graduation rates initially stagnated or restarted, and significant gaps based on race, gender, and socioeconomic status still exist. The role of HSEEs in these trends is also unknown.

The impact of HSEEs on graduation rates remains difficult to determine, especially in the long term. This is especially difficult because of the lack of controlled experimental conditions; counterfactual conditions and randomization simply do not exist. Regression discontinuity designs around the margin of barely passing have been very useful in identifying the effects of exams for students' trajectories and graduation probabilities (Dee and Jacob, 2006; Reardon, Arshan, Atteberry, and Kurlaender, 2010), but overall effects are difficult to determine. We examine long-term trends in graduation rates surrounding and following the year in which HSEEs were first used to withhold diplomas. By by exploiting the temporal variation in state-level implementation of HSEEs, we are able to at least partially mitigate year- and state-specific trends.

4 Data and Descriptive Statistics

In this section, we briefly describe our data collection process and present descriptive statistics. We combine data from multiple sources into a unique data set. The main source of information is the Center on Education Policy (CEP), which gathers state reports on high school education and high school examination procedures. From CEP's state reports we know whether a given state has an HSEE, when it was first administered (or reformed), and how it is structured in terms of grade alignment and content.[1]

To construct a panel, we complemented the data from CEP with information on high school graduation rates and achievement scores. We took graduation rates for the period 1990-2012 from the Digest of Education Statistics (2012), produced by the National Center for Educational Statistics (NCES). We completed the NCES graduation rate data with information from America's Health Ranking,[2] which has prepared annual reports on health and health dynamics since 1990 that include high school graduation rates for each state from 1990 to 2013. For achievement, the other outcome of our interest, we rely on the NCES, which is also responsible for gathering data on the National Assessment of Educational Progress (NAEP). The NAEP is the most representative long-term assessment of the skills and abilities of American students, and is reported at the state level in 4^{th} and 8^{th} grades. Assessments are not conducted every year, but still represent the best source for comparable state-level achievement data; we use 8^{th} grade math scores as our measure of achievement. NAEP scores are a well-known measure of achievement in the research community.[3] Altogether, the resulting data forms a longitudinal state-level aggregated data set covering the period 1990-2013. The panel is balanced for the outcome "graduation rate" but not for the outcome "achievement," because NAEP scores are not collected every year.

Table 1 presents descriptive statistics, shown separately for all states (Panel A) and only for those that introduced an HSEE (Panel B). For a detailed table with HSEE introduction years for each state, see appendix Table A.1. Table 1 shows that 56 percent of states—28 in total—have introduced an HSEE.[4] Almost all states that have an HSEE introduced it between 1990 and 2012, excepting Alabama, New York, and South Carolina.

[1] We also double-checked our information with that of Dee and Jacob (2006), finding almost no difference.

[2] http://www.americashealthrankings.org/

[3] See Dee, Evans, and Murray (1999)

[4] We exclude the District of Columbia from our entire analysis.

Table 1: Descriptive Statistics

	Mean [1]	Std. Dev. [2]	Min [3]	Max [4]
Panel A. All States (50 States, $N = 1,200$)				
Year	2001.50	6.93	1990	2013
HSEE	0.56	0.50	0.0	1.0
Graduation Rate	73.89	8.40	48.0	91.6
NAEP Score (8^{th} Grade Math)	277.02	10.30	246.0	301.0
Panel B. States with HSEE (28 States, $N = 672$)				
Year	2001.50	6.93	1990	2013
First HSEE Administered	1999.86	5.53	1990	2012
Graduation Rate	70.24	8.31	48.0	89.7
NAEP Score ($^{\text{th}}$ Grade Math)	274.54	10.85	246.0	301.0

Notes: Data collected by the authors.

These three states had an HSEE before 1990,[5] but reformed it in the 1990-2012 period. In our econometric analysis, we consider the reform year as the year of HSEE introduction for these three states.

Table 1 also presents descriptive statistics for our outcomes of interest. Graduation rates are a topic of much discussion in the HSEE literature because they are the area where their effects are most obvious. Most of this literature focuses on individual states in the short term following the application of an exam; we investigate graduation rates over the longer term and across all states with HSEEs. The average graduation rate for all states in the period 1990-2013 is about 74 percent, with a minimum of 48 percent in South Carolina (in 2003) and a maximum of 91.6 in Vermont (in 1992). The average graduation rate for the states that introduced an HSEE is almost four percentage points below the national average, which begs the question of whether their introducing an HSEE somehow helped the states without an HSEE to catch up; we answer this question in our robustness checks.

The effects of HSEEs for achievement and the quality of education is a more challenging question, especially for graduates' achievement following high school. Part of the

[5] Alabama introduced an HSEE in 1984, New York in 1878 (Regents examination), and South Carolina in 1986.

intended effect of HSEEs on education quality is their ability to focus student, teacher, and school effort on ensuring students' knowledge of minimum requirements, so we assume that the implementation of an HSEE would trigger system-wide efforts to increase attainment. For this reason and to avoid any potential effects of dropout in 12$^{\text{th}}$ grade scores, we use 8th grade NAEP mathematics scores to represent the level of achievement in each state's educational system as a whole. We choose mathematics because state standards are clearest and most nationally consistent on that subject. The national average NAEP score for 8$^{\text{th}}$ grade math is 277, whereas the average for states that introduced an HSEE is 274.5. Because NAEP scores are not collected annually, we only have 457 observations for this outcome (260 for the sub-sample of states with an HSEE). This limits the potential significance of our results but does not compromise the integrity of our analysis as an indication of trends in achievement.

5 Empirical Strategy

State-level time trends in high school graduation rates and achievement are a natural point of departure for considering the impact of introducing an HSEE on these two outcomes. Figure 1 and Figure 2 present trends in state-level graduation rates and eighth grade NAEP math scores from 1990 to 2013. The vertical line visually identifies the relative point in time when an HSEE was implemented. The trends shown in Figures 1 and 2 suggest that introducing an HSEE may have had some positive effects on graduation rates and eighth-grade math scores, at least for those states that did introduce one. However, potential state- or nationwide changes in educational, social, and economic factors over this period make it difficult to credibly identify causal inferences from these trends.

To circumvent these concerns, we use an interrupted time series (ITS) approach. The basic intuition of ITS designs is to use pre-treatment trends in the outcome as a

Figure 1: Average Graduation Rate over Time Relative to HSEE Introduction (28 States)

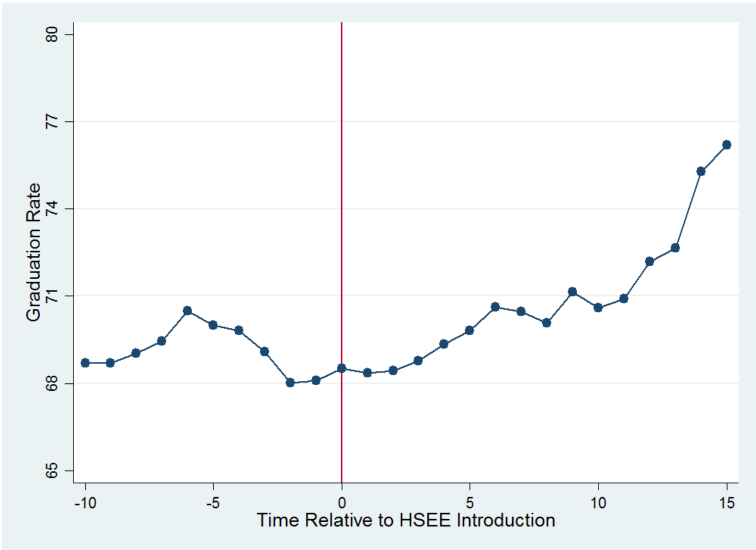
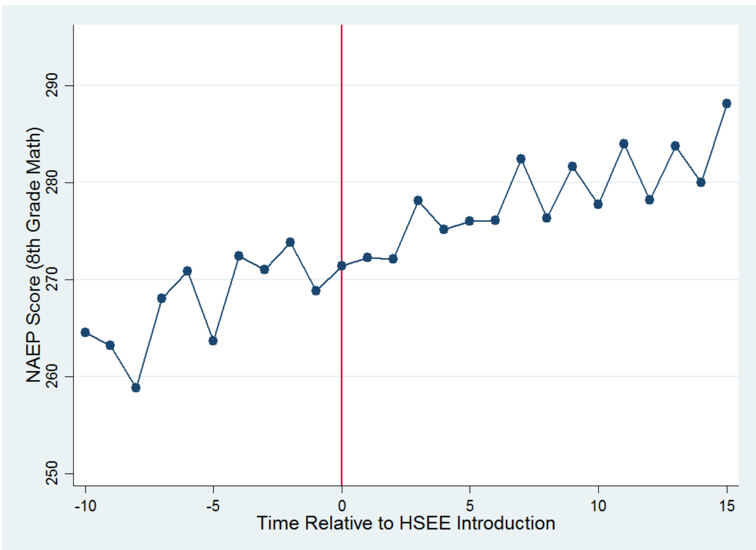


Figure 2: Average Achievement over Time Relative to HSEE Introduction (28 States)



counterfactual—what the post-treatment trend in the outcome would have been in the absence of the treatment. Given that this assumption might be overly strong in some cases—especially if other factors have changed together with the treatment—we perform a robustness check using states that do not have an HSEE as control group. The ITS approach has a long tradition in education research,[6] and has been used recently to evaluate several educational reforms such as No Child Left Behind (Dee and Jacob, 2011; Dee, Jacob, and Schwartz, 2012) and Accelerated Schools (Bloom, Ham, Melton, and O'Brien, 2001).

In general, we can describe the trend of an outcome Y in state i at time t as:

$$Y_{it} = f(t - t_i^*) + HSEE_{it} \cdot g(t - t_i^*) + \gamma_i + \varepsilon_{it} \tag{1}$$

where $HSEE_{it}$ indicates the treatment status of unit i at time t, which in our case is an indicator of whether the state has an HSEE at time t. Time is measured in both continuous (t) and relative (t_i^*) metrics. Specifically, t_i^* denotes the time at which treatment begins in unit i; therefore t_i^* is the period such that $HSEE_{it} = 1$ if $t \geq t_i^*$ and $HSEE_{it} = 0$ if $t < t_i^*$. Subscript i in in the relative metric of time is necessary because states introduced their HSEEs at different points in time. This strengthens our causal claims by filtering out other changes or policies that may have happened concurrently. The function f describes the trend prior to t_i^*, and starting from period t_i^* the trend in Y_{it} is described by the function $(f + g)$. Under the assumption that the trend described by f would have continued after t_i^* in the absence of treatment, the effect of $HSEE_{it}$ by time $t \geq t_i^*$ is given by $g(t - t_i^*)$. Finally, γ_i represents state fixed effects and ε_{it} is an error term.

In ITS designs, multiple effects exist. First, we can estimate a sharp discontinuity at the time of intervention—a change in level. Second, we are able to estimate the change

[6] For an overview, see Shadish, Cook, and Campbell (2002).

in the slope of the time-series at the point of intervention—a kink point. Third, we can estimate a continuous (or discontinuous) effect that does not (or does) decay over time. Fourth, it is possible to study potential intervention effects that are immediate, delayed, or even anticipatory. Depending on the functional form we impose on the functions f and g in Equation 1, we can focus on any effects of interest. Most simply, we might approximate f and g as linear functions of time as follows:

$$Y_{it} = \beta_0 + \beta_1 \cdot (t - t_i^*) + \beta_2 \cdot HSEE_{it} + \beta_3 \cdot (t - t_i^*) \cdot HSEE_{it} + \gamma_i + \varepsilon_{it} \tag{2}$$

The model specified in Equation 2 says that the trend in Y_{it} before time t_i^* is linear with slope β_1. At time t_i^*, the value of Y_{it} changes by β_2, then the trend in Y_{it} after time t_i^* is linear with slope $(\beta_1 + \beta_3)$. The effect of introducing an HSEE by time $t \geq t_i^*$ is given by $\beta_2 + \beta_3 \cdot (t - t_i^*)$. We can test the null hypothesis that the effect at time t is zero with the following F-test: $\beta_2 + \beta_3 \cdot (t - t_i^*) = 0$. In our linear specification, β_2 can be seen as a regression discontinuity estimate of the immediate effect of $HSEE_{it}$ on Y_{it}, with the difference that in ITS we observe one unit at different points in time whereas in a regression-discontinuity design we would observe multiple units at one point in time. Similarly, β_3 can be seen as a difference-in-differences estimate of the effect of $HSEE_{it}$ on Y_{it}, or the difference in the rate of change for the average Y_{it} between treated and untreated states.

As we explained, Equation 2 allows us to estimate two effects, namely the change in level and the change in slope at the time of intervention. However, the treatment effect might decay or reinforce itself over time, and in cases like these our linear specification would lose the pattern. To allow the effects of the introduction of an HSEE to be a non-parametric function of time, we specify the following model:

$$Y_{it} = \beta_0 + \beta_1 \cdot (t - t_i^*) + \sum_{j=0}^{J} \delta_j \cdot D_t^j + \gamma_i + \varepsilon_{it} \tag{3}$$

where is a dummy variable equal to one if $(t - t_i^*) = j$, and J is the number of years observed after t_i^* (in our case $J = 15$). In Equation 3, the effect of $HSEE_{it}$ on Y_{it} j years after the introduction of an HSEE is represented by δ_j, and we can test the null hypothesis of no treatment effect at year j by testing $H_0 : \delta_j = 0$.

As we discussed before, another interesting effect that can be seen in in ITS models is the anticipation effect, or whether the intervention has an effect on the outcome before it is actually introduced. To find this, we allow the effect of introducing an HSEE to be a fully non-parametric function of time:

$$Y_{it} = \beta_0 + \sum_{j=-10}^{J} \delta_j \cdot D_t^j + \gamma_i + \varepsilon_{it} \tag{4}$$

where D_t^j is, again, a dummy variable equal one if $(t - t_i^*) = j$, and J is the number of years observed after t_i^* ($J = 15$). In this specification, j can also be negative to estimate the trend in Y_{it} for each pre-treatment period up to ten years before the introduction of an HSEE. The non-parametric approach we use in Equation 4 is well known and used among labor economist, for example in estimating wage losses after job separation.[7]

In time-series settings, estimating pre-intervention effects non-parametrically also constitutes an important test for causality. As suggested in the early literature by Granger (1969, 1988), the availability of many years of data enables testing of whether changes in a given policy lead or lag the outcome. If the within-state changes in Y_{it} lag—or coincide with—the within-state intervention, then they are consistent with a causal story in which causes are followed by their effects. In contrast, if within-state changes in the outcome lead the within-state changes in the policy, we might normally suspect either policy endogeneity or the presence of unobserved time-varying state characteristics that drive the effects. In our setting, however, it is hard to believe that introducing an HSEE had

[7]See, for example, Jacobson, LaLonde, and Sullivan (1993); von Wachter, Song, and Manchester (2008); Balestra and Backes-Gellner (2012).

no pre-intervention effects, especially because such a policy follows a lengthy political discussion involving schools, districts, and the state government. Therefore, we might expect an effect of the HSEE on graduation rates or even achievement a few years before its actual introduction. For example, if schools know that an exit exam is going to be introduced soon, they might modify their current high school curricula in anticipation to ensure that students are prepared.

In sum, to estimate the effect of introducing an HSEE on graduation rates and achievement, we use an ITS approach and three different specifications (Equations 2, 3 and 4). We rely on different specifications because we are interested in the many effects an HSEE might have. While the linear specification (Equation 2) focuses on the changes in level and slope at the time of intervention, the semi-parametric (Equation 3) and fully non-parametric (Equation 4) specifications allow the estimation of treatment effects specific for each year before, during, and after the intervention.

6 Results

This section presents our results in three parts. The first subsection shows the results for the graduation rate outcome, starting from the linear specification (Equation 2) and then presenting our semi-parametric (Equation 3) and non-parametric specifications (Equation 4). The second subsection examines the effect of introducing an HSEE on achievement. In the third subsection, we perform robustness checks to test the internal and external validity our empirical strategy.

6.1 Effect of HSEE on Graduation Rates

Table 2 presents regression outputs for the linear specification in Equation 2. We estimate three effects: the trend in graduation rates before the introduction of an HSEE $(t - t_i^*)$,

Table 2: Effect of HSEE on Graduation Rate, Linear Specification

Variables	Graduation Rate	
	Coefficient	Standard Error
	[1]	[2]
$(t - t_i^*)$	-0.141	(0.118)
HSEE	-0.871	(1.129)
$(t - t_i^*)\cdot$ HSEE	0.562***	(0.144)
Intercept	68.447***	(0.885)
State Fixed Effects	YES	
Adjusted ^{2}	0.777	
N	672	

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Standard errors are clustered at the State level and robust to heteroskedasticity and serial correlation.

the change in level of graduation rates in the year of introduction ($HSEE$), and the change in slope after the introduction of an HSEE (the interaction term). Note that we are estimating a within-state effect for the states that introduced an HSEE, thus the counterfactual is the state's pre-intervention trend in graduation rates.[8]

From Table 2 we infer that there is no particular trend in graduation rates before the introduction of an HSEE, because the coefficient of $(t - t_i^*)$ is not significant. Similarly, there is no significant discontinuity at the time of intervention. However, and most important for our research question, we estimate a positive change in slope for the time-series at the time of intervention. This kink point is significant at the highest confidence level (p-value $= 0.00$), and means that introducing an HSEE has a positive impact on the trend in graduation rates.

In sum, according to our linear specification, the within-state graduation rate trend before introducing an HSEE is rather stable, then becomes positive after implementation. This positive pattern is consistent with previous research on exit exams for industrialized countries (Bishop, Mane, and Bishop, 2001). However, by specifying a linear function of time we are assuming that the effect of introducing an HSEE is constant and are

[8]We consider alternative counterfactuals in subsection 5.3, in which we perform our robustness checks.

not allowing for heterogeneous effects over time. We might suspect that this assumption invalid if the immediate effect decays or reinforces over time. Therefore, we also estimated semi- and non-parametric models.

Table 3 shows the results of the semi-parametric (columns 1-2) and non-parametric (columns 3-4) specifications. The semi-parametric model assumes a linear trend in graduation rates before HSEE introduction and a distinct effect in each year thereafter. Although we know from Table 2 that the overall trend in graduation rates becomes positive after introducing an HSEE, the semi-parametric specification reveals that the trend is initially negative. For the first four years with a new HSEE, within-state graduation rates decrease by almost three percentage points. This loss is statistically significant and negative in the first four years, at which point it becomes insignificant.

A similar picture emerges from the non-parametric specification, which estimates the time-series as a non-parametric function of time with the treatment effect estimated separately for each year before and after HSEE introduction. In the third column of Table 3, we observe that the significant short-term loss lasts up to three years after introducing an HSEE and has a magnitude of slightly more than three percentage points per year. Note that in the non-parametric specification the estimated effects are relative to the year of introduction.[9] In the long term, the effect of introducing an HSEE even becomes positive and marginally significant, which drives the results in the linear specification.

In the non-parametric specification, we can study not only effects following the introduction of an HSEE but also the potential anticipatory effects. As discussed in Section 4, we might expect states to adjust their curricula before the actual introduction of an HSEE in order to prepare their students, teachers, and schools for the new HSEE requirement. Furthermore, major educational policy changes like HSEEs are usually accompanied by several years of political and popular discussion and we can easily assume

[9] Changing the base category (e.g., setting the first year as base category) has no impact on the results.

Table 3: EFFECT OF HSEE ON GRADUATION RATE, NON-PARAMETRIC SPECIFICATIONS

Variables	Graduation Rate			
	Coefficient	Standard Error	Coefficient	Standard Error
	[1]	[2]	[3]	[4]
$(t - t_i^*)$	0.198*	(0.072)		
$(t_i^* - 10)$			-1.710	(1.145)
$(t_i^* - 9)$			-2.020	(1.162)
$(t_i^* - 8)$			-2.045	(1.181)
$(t_i^* - 7)$			-1.793	(1.246)
$(t_i^* - 6)$			-1.433	(1.251)
$(t_i^* - 5)$			-1.638	(1.183)
$(t_i^* - 4)$			-1.834	(1.289)
$(t_i^* - 3)$			-2.551	(1.257)
$(t_i^* - 2)$			-3.617*	(1.337)
$(t_i^* - 1)$			-3.644**	(1.087)
(t_i^*)	-2.448**	(0.663)	*Base Category*	
$(t_i^* + 1)$	-2.785**	(0.774)	-3.379**	(1.040)
$(t_i^* + 2)$	-2.723**	(0.937)	-3.193**	(1.100)
$(t_i^* + 3)$	-2.732*	(1.084)	-3.076*	(1.217)
$(t_i^* + 4)$	-2.342	(1.276)	-2.487	(1.451)
$(t_i^* + 5)$	-1.976	(1.235)	-1.985	(1.364)
$(t_i^* + 6)$	-1.374	(1.410)	-1.185	(1.545)
$(t_i^* + 7)$	-1.712	(1.520)	-1.325	(1.665)
$(t_i^* + 8)$	-2.371	(1.655)	-1.828	(1.768)
$(t_i^* + 9)$	-1.988	(1.814)	-1.285	(2.074)
$(t_i^* + 10)$	-2.271	(1.623)	-1.405	(1.814)
$(t_i^* + 11)$	-2.165	(1.770)	-1.101	(1.902)
$(t_i^* + 12)$	-1.068	(1.978)	0.195	(2.058)
$(t_i^* + 13)$	-0.758	(1.847)	0.547	(1.823)
$(t_i^* + 14)$	0.330	(1.248)	1.747	(1.231)
$(t_i^* + 15)$	1.064	(1.302)	2.606	(1.304)
Intercept	70.951***	(0.515)	71.743***	(0.946)
State FE	YES		YES	
Adjusted R^2	0.760		0.749	
N	672		672	

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Standard errors are clustered at the State level and robust to heteroskedasticity and serial correlation.

that students, teachers, and schools were aware of the impending introduction of an exam. The results in column 3 of Table 3 confirm our expectations, with statistically significant pre-intervention effects one and two years before HSEE introduction. The presence of pre-intervention effects, however, might weaken our causal claims because it might appear that the effects are leading—rather than lagging—the cause. Still, the availability of many years of data and the fact that stakeholders are aware of HSEEs before their introduction reinforces our interpretation of the results.

In sum, we find that introducing an HSEE has an overall positive effect on graduation rates and a positive effect on the slope of the time-series for graduation rates. However, this effect on graduation rates is heterogeneous over time. In the year of introduction and for at least the following three years, HSEEs have a negative impact on graduation rates. This negative impact is short-lived and becomes positive towards the end of our time span. We also estimate a pre-intervention negative effect one and two years before HSEE introduction, suggesting that students, teachers, and schools start preparing for exams even before their actual introduction.

6.2 Effect of HSEE on Achievement

Our second outcome of interest is achievement. The only measure of achievement we have for all states and for multiple years is the NAEP score. As many other studies do,[10] we rely on 8^{th} grade math scores as a measure for achievement. We choose math because mathematical skills are relatively easier to measure through standardized tests, state standards for mathematics are most consistent, and the potential for bias is minimized. Similarly, we choose 8^{th} grade because we want to analyze a grade as close as possible to high school without introducing the possibility of selection bias from dropout.

We believe that a introducing an HSEE impacts the behavior of students and teachers

[10] See, for example, Dee and Jacob (2011)

Table 4: Effect of HSEE on Achievement, Linear Specification

Variables	NAEP Scores (8^{th} Grade Math) Coefficient [1]	Standard Error [2]
$(t - t_i^*)$	0.912***	(0.159)
HSEE	0.627	(1.009)
$(t - t_i^*) \cdot$ HSEE	0.083	(0.180)
Intercept	270.562***	(0.952)
State Fixed Effects	YES	
Adjusted ^{2}	0.927	
N	260	

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Standard errors are clustered at the State level and robust to heteroskedasticity and serial correlation.

and the development and application of curriculum not only at the high school level but throughout a state's education system. The general trend of education policy in the United States over the past decades supports this intuition: HSEEs themselves are part of a broader trend towards streamlining, quantifying, and building accountability in education. As is especially obvious following the introduction of the Common Core State Standards, states are attempting to set and assess clear and rigorous standards at all grade levels.

Table 4 presents regression outputs for the linear specification. The number of observations is halved because we have an unbalanced panel for the achievement outcome. The only significant coefficient in Table 4—aside from the intercept—is the linear trend in time, which is positive and highly significant. Therefore, it appears that introducing an HSEE has no impact on the (positive) trend in achievement: we estimate neither a discontinuity nor a kink.

Table 5 shows estimated effects of HSEEs on achievement for the alternative semi- and non-parametric specifications. Similarly to Table 4, we estimate no particular effect of introducing an HSEE on achievement. We nevertheless observe a positive trend in our non-parametric specification—a trend that becomes statistically significant about ten years after HSEE introduction. Note that the multitude of non-significant coefficients in the non-parametric specification is likely due to the relatively small sample size and the many degrees of freedom we lose estimating all the parameters.

We conclude that introducing an HSEE has no statistically significant impact on the time-series of achievement. This result holds for the states that introduced an HSEE, which have lower achievement on average compared to those that never had an HSEE. In the next subsection we perform the analysis including states that have no HSEE as part of the control group. This not only constitutes a robustness check for our results, but also shows whether introducing an HSEE helps to close the achievement gap between states with HSEEs and states without by raising the level of the states that introduced HSEEs from much lower than their non-HSEE counterparts to only slightly lower.

6.3 Robustness Checks

In this subsection, we estimate our models for the full set of all states. Doing so allows us to use states without HSEEs as counterfactuals for those that have one. Including all of the states in the regressions has two purposes. First, we test the robustness of our results. If we find completely different effects, it would mean that states with HSEEs are completely different from those without HSEEs, casting some doubt on the external validity of our results. However, if the results are in fact consistent with those of the previous subsections, it would mean that states are rather similar among themselves—at least after controlling for time-invariant characteristics. This might increase the generalizability of our results to all states. Second, using states without HSEEs as counterfactuals might have relevant policy implications. Given that states with HSEEs have lower graduation rates and lower achievement, we can investigate whether introducing an HSEE helps close the graduation rate and achievement gaps between states.

Table 6 shows regression outputs for the effect of HSEEs on graduation rate (columns

Table 5: Effect of HSEE on Achievement, Non-Parametric Specifications

Variables	NAEP Scores (8^{th} Grade Math)			
	Coefficient	Standard Error	Coefficient	Standard Error
	[1]	[2]	[3]	[4]
$(t - t_i^*)$	0.994***	(0.066)		
$(t_i^* - 10)$			-9.300*	(4.111)
$(t_i^* - 9)$			-7.257	(3.720)
$(t_i^* - 8)$			-8.502*	(3.796)
$(t_i^* - 7)$			-5.884	(4.620)
$(t_i^* - 6)$			-5.366	(4.034)
$(t_i^* - 5)$			-5.266	(3.775)
$(t_i^* - 4)$			-0.776	(3.202)
$(t_i^* - 3)$			-0.650	(5.150)
$(t_i^* - 2)$			-0.874	(3.849)
$(t_i^* - 1)$			-2.563	(4.626)
(t_i^*)	-0.822	(1.730)	*Base Category*	
$(t_i^* + 1)$	0.510	(1.001)	-0.198	(5.914)
$(t_i^* + 2)$	-1.187	(1.025)	0.321	(4.041)
$(t_i^* + 3)$	2.636*	(0.979)	5.981	(3.624)
$(t_i^* + 4)$	-0.754	(1.141)	3.465	(3.512)
$(t_i^* + 5)$	1.319	(0.916)	5.169	(4.330)
$(t_i^* + 6)$	0.520	(1.160)	5.577	(4.195)
$(t_i^* + 7)$	1.513	(1.004)	8.463*	(3.490)
$(t_i^* + 8)$	-0.368	(1.203)	7.050	(3.961)
$(t_i^* + 9)$	0.978	(1.639)	8.611	(4.283)
$(t_i^* + 10)$	-0.054	(1.162)	7.836	(4.200)
$(t_i^* + 11)$	0.680	(1.740)	10.809**	(3.700)
$(t_i^* + 12)$	-1.408	(1.145)	8.402	(4.319)
$(t_i^* + 13)$	-0.221	(1.576)	10.004*	(3.945)
$(t_i^* + 14)$	-1.778	(1.094)	7.412	(4.268)
$(t_i^* + 15)$	0.234	(1.182)	11.218**	(4.038)
Intercept	271.185***	(0.412)	272.067***	(3.036)
State FE	YES		YES	
Adjusted R^2	0.927		0.636	
N	260		260	

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Standard errors are clustered at the State level and robust to heteroskedasticity and serial correlation.

Table 6: Effect of HSEE on Graduation Rate and Achievement, All States

Variables	Graduation Rate		NAEP Scores (8^{th} Grade Math)	
	Coefficient	Standard Error	Coefficient	Standard Error
	[1]	[2]	[3]	[4]
$(t - t_i^*)$	0.033	(0.045)	0.752***	(0.061)
HSEE	-1.946*	(0.936)	1.670	(0.903)
$(t - t_i^*)\cdot$ HSEE	0.395***	(0.094)	0.234*	(0.100)
Intercept	73.335***	(0.296)	269.791***	(0.348)
State Fixed Effects	Yes		Yes	
Adjusted R^2	0.799		0.921	
N	1,200		457	

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Standard errors are clustered at the State level and robust to heteroskedasticity and serial correlation.

1-2) and achievement (columns 3-4). We restrict our analysis to the linear specification for easy comprehension and because the semi- and non-parametric models—which are presented in appendix Tables B.1 and B.2—reveal no additional information. First and foremost, we observe that the effects are very similar to those of the previous tables, especially in terms of direction. However, we find some differences in terms of significance, which might be partly due to the increase in sample size.

One relevant finding of Table 6 is that introducing an HSEE appears to help close both graduation rate and achievement gaps between states with HSEEs and those without. For graduation rates, we estimate a negative and significant discontinuity at the time of HSEE introduction of about two percentage points. We also estimate a highly significant change in the slope of the graduation rate time-series. Regarding achievement, we estimate the same highly significant and positive long-term trend as before. Moreover, we also find a positive and significant change in slope for those states that introduced an HSEE after its introduction. HSEEs alone may not cause these increasingly positive trends in graduation rates and achievement, but some part of the streamlining and focusing of educational standards that includes HSEEs and policy does cause them.

Overall, we conclude that our results are not sensitive to the control group chosen. Additionally, we find that introducing an HSEE helps reduce the graduation rate and achievement gap between states with HSEEs (usually lower performing) and states without HSEEs (usually higher performing).

7 Conclusions and Discussion

There are some potential mechanisms behind the improvement in graduation rate trends and—in some specifications—the similar improvement in achievement growth following the implementation of an HSEE. The first possibility is improvement in curriculum and the behavior of schools, teachers, and students. This is the intended effect of HSEEs, and the increased accountability from the exam could cause such improvements. Second, the focusing of curriculum, instruction, and student effort could be responding to the perverse narrowing incentive to only the material covered on the exam. In the case of an HSEE that perfectly represents and measures the skills and knowledge of a high school diploma, this would be identical to the first possibility. If there were design or measurement issues in the exam, however, this would place limitations on the value and effectiveness of students' education. In both of these cases, the improvement in graduation rates and math test achievement we find would be explained by the HSEE successfully modifying educational behavior towards increased achievement—at least as measured by the exam itself.

Alternatively, changes in the HSEE such that the exam adapts to a state's students rather than the other way around could yield the increase in graduation rates observed in this study; instead of the HSEE modifying behavior as intended, revisions of the exam due to political or social pressure may improve students' likelihood of passing. One possibility for such an adjustment—one that has been called for and implemented frequently—is an option for non-native English speakers to take the exam in their native language. HSEE policies have been accompanied by an outcry against potential discrimination against immigrant students in many states, and most of those states have added concessions for non-English speakers. These concessions increase the passing rate for the groups of students they target, which would raise the overall passing rate. A second possible adjustment would be to simply lower the HSEE's difficulty level or adjust its format in response to claims that it was too rigorous or discriminatory. Many states are moving towards end-of-course rather than comprehensive exams, and the content of exams changes frequently. These changes may also contribute to higher passage rates on average.

Over the long term, HSEEs do not appear to decrease graduation rates, instead they strengthen improvement trends in graduation rates. More importantly, this increase in graduation rates is not accompanied by any decrease in achievement—more students are graduating and meeting the increased standard. Graduation dips in the short term immediately before, during, and after the year in which exam scores are first used to withhold diplomas, but recovers soon after and even improves over pre-HSEE trends. The mechanism is still unclear; while improvements may come from the adjustments in school, teacher, and student behavior as intended by policymakers, improvements in graduation rates may also come from narrowing of teaching and learning to exam material or changes in the exam itself rather than student achievement. Further research on how HSEEs interact with graduation rates and especially achievement is necessary, but the first step has been to understand how they affect graduation rates and achievement overall.

It is true that students at the margin of failing the exam whose graduation year is on or near the first year of an exit exam policy will have decreased probabilities of graduation, but these effects are not persistent and are counterbalanced by later improvements in student attainment. If exams succeed in modifying school, teacher, and student behavior towards a more focused mastery of the material required to graduate high school, the exams are useful insofar as they accurately reflect and measure such material. These findings may help rectify the lack of convergence in the literature towards a standard result for HSEEs and graduation rates; the net effect of HSEEs for graduation rates is positive over the long term, but shorter-term effects can be negative. By estimating both effects, we rectify some of the disagreement in the literature emerging from differences in data sources and estimation strategies. The simultaneous effects of raising attainment by enhancing behavior and improving the signaling value of a high school diploma are important steps forward for secondary education.

References

Akerlof, G. A. (1970): "The market for "lemons": Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, 84(3), 488–500.

Amrein, A. L., and D. C. Berliner (2002): "High-stakes testing & student learning," *Education Policy Analysis Archives*, 10, 1–74.

Arcidiacono, P., P. Bayer, and A. Hizmo (2010): "Beyond signaling and human capital: Education and the revelation of ability," *American Economic Journal: Applied Economics*, 2(4), 76–104.

Balestra, S., and U. Backes-Gellner (2012): "When a door closes a window opens? Investigating the effects of involuntary separations," Economics of Education Working Paper 0072, University of Zurich, Institute for Strategy and Business Economics (ISU).

Bishop, J. H. (1997): "The effect of national standards and curriculum-based exams on achievement," *American Economic Review*, 80(2), 260–264.

Bishop, J. H., and F. Mane (2001): "The impacts of minimum competency exam graduation requirements on high school graduation, college attendance and early labor market success," *Labour Economics*, 8(2), 203–222.

Bishop, J. H., F. Mane, and M. Bishop (2001): "How external exit exams spur achievement," *Educational Leadership*, 59(1), 58–63.

Bitler, M. P., J. B. Gelbach, and H. W. Hoynes (2006): "What mean impacts miss: Distributional effects of welfare reform experiments," *American Economic Review*, 96(4), 988–1012.

Bloom, H. S., S. Ham, L. Melton, and J. O'Brien (2001): "Evaluating the Accelerated Schools Approach: A Look at Early Implementation and Impacts on Student Achievement in Eight Elementary Schools.," Evaluative Report 142, Manpower Demonstration Research Corporation.

Booher-Jennings, J. (2005): "Below the bubble: "Educational triage" and the Texas accountability system," *American Educational Research Journal*, 42(2), 231–268.

Carnoy, M., and S. Loeb (2002): "Does external accountability affect student outcomes? A cross-state analysis," *Educational Evaluation and Policy Analysis*, 24(4), 305–331.

Dee, T. S., W. N. Evans, and S. E. Murray (1999): "Data watch: Research data in the economics of education," *Journal of Economic Perspectives*, 13(3), 205–216.

Dee, T. S., and B. Jacob (2011): "The impact of no child left behind on student achievement," *Journal of Policy Analysis and Management*, 30(3), 418–446.

Dee, T. S., B. Jacob, and N. L. Schwartz (2012): "The effects of NCLB on school resources and practices," *Educational Evaluation and Policy Analysis*, 35(2), 252–279.

Dee, T. S., and B. A. Jacob (2006): "Do high school exit exams influence educational attainment or labor market performance?," NBER Working Paper 12199, National Bureau of Economic Research, Inc.

Frisancho, V., K. Krishna, S. Lychagin, and C. Yavas (2013): "Better luck next time: Learning through retaking," IDB Working Paper 483, Inter-American Development Bank.

Gibbons, R., and L. F. Katz (1991): "Layoffs and lemons," *Journal of Labor Economics*, 9(4), 351–380.

Granger, C. W. J. (1969): "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, 37(3), 424–438.

Granger, C. W. J. (1988): "Some recent development in a concept of causality," *Journal of Econometrics*, 39(1-2), 199–211.

Greene, J. P., and M. A. Winters (2005): "Public high school graduation and college-readiness rates: 1991-2002," Education Working Paper 8, Center for Civic Innovation.

Grodsky, E., J. R. Warren, and D. Kalogrides (2009): "State high school exit examinations and NAEP long-term trends in reading and mathematics, 1971-2004," *Educational Policy*, 23(4), 589–614.

Jackson, E., and M. E. Page (2013): "Estimating the distributional effects of education reforms: A look at Project STAR," *Economics of Education Review*, 32(C), 92–103.

Jacob, B. A. (2001): "Getting tough? The impact of high school graduation exams," *Educational Evaluation and Policy Analysis*, 23(2), 99–121.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993): "Earnings losses of displaced workers," *American Economic Review*, 83(4), 685–709.

Marchant, G. J., and S. E. Paulson (2005): "The relationship of high school graduation exams to graduation rates and SAT scores.," *Education Policy Analysis Archives*, 13(6), 1–17.

Murnane, R. J. (2013): "U.S. high schoolig graduation rates: Patterns and explanations," *Journal of Economic Literature*, 51(2), 370–422.

Ou, D. (2010): "To leave or not to leave? A regression discontinuity analysis of the impact of failing the high school exit exam," *Economics of Education Review*, 29(2), 171–186.

Plunk, A. D., W. F. Tate, L. J. Bierut, and R. A. Grucza (2014): "Intended and unintended effects of state-mandated high school science and mathematics course graduation requirements on educational attainment," *Educational Researcher*, 43(5), 230–241.

Reardon, S. F., N. Arshan, A. Atteberry, and M. Kurlaender (2010): "Effects of failing a high school exit exam on course taking, achievement, persistence, and graduation," *Educational Evaluation and Policy Analysis*, 32(4), 498–520.

Shadish, W. R., T. D. Cook, and D. T. Campbell (2002): *Experimental and quasi-experimental designs for generalized causal inference.* Wadsworth Cengage Learning.

Spence, A. M. (1973): "Job market signaling," *The Quarterly Journal of Economics*, 87(3), 355–374.

Tyler, J. H., R. J. Murnane, and J. B. Willet (2000): "Estimating the labor market signaling value of the GED," *Quarterly Journal of Economics*, 115(2), 431–468.

von Wachter, T., J. Song, and J. Manchester (2008): "Long-term earnings losses due to job separation during the 1982 recession: An analysis using longitudinal administrative data from 1974 to 2004," Discussion paper.

Warren, J. R., and M. R. Edwards (2005): "High school exit examinations and high school completion: Evidence from the early 1990s," *Educational Evaluation and Policy Analysis*, 27(1), 53–74.

Warren, J. R., K. N. Jenkins, and R. B. Kulick (2006): "High school exit examinations and state-level completion and GED rates, 1975 through 2002," *Educational Evaluation and Policy Analysis*, 28(2), 131–152.

APPENDIX

A High School Exit Exam per State

Table A.1: List of States with an HSEE and Their Year of First Administration

State	Type of Test	Year(s) first Administered	Year Diplomas First Withheld	Grade First Administered	Grade(s) Exam Aligned to
Alabama	SB	1984, 1995	1985, 2001	10	11
Alaska	SB	2000	2004	10	8 to 10
Arizona	SB	1999	2006	10	10
Arkansas	EOC	2001, 2010	N/A	Algebra 1	Algebra 1
California	SB	2001, 2004	2006	10	10, Algebra 1
Florida	SB	1998	2003	10	10
Georgia	SB	1991	1994	11	9 to 11
Idaho	SB	2004	2006	10	10
Indiana	SB	1997	2000	10	9
Louisiana	SB	2001	2003	10	9 to 12
Maryland	EOC	2001	1989, 2009	DOS	10
Massachusetts	SB	1998	2003	10	10
Minnesota	SB	1996, 2010	2000, 2010	9 to 11	8 to 10
Mississippi	EOC	2000, 2007	2006	DOS	9 to 11
Nevada	SB	2001	2003	10	9 to 12
New Jersey	SB	1991, 2002	2003	11	11
New Mexico	MC	2011	2012	11	9 to 12
New York	RE	1878, 2000	2003	DOS	9 to 12
N. Carolina	EOC	2006	2010	DOS	Course-specific
Ohio	SB	1990, 2005	1994, 2007	10	10
Oklahoma	EOC	2001	2012	DOS	HS standards
Oregon	SB	2009	2012	3	11
Rhode Island	CO	2012	2012	11	9, 10
S. Carolina	SB	1986, 2005	2006	9	9
Tennessee	EOC	2001	2005	DOS	10
Texas	SB	1990, 2003	2005	11	HS standards
Virginia	EOC	1998	2004	DOS	Course-specific
Washington	SB	1999, 2010	2008, 2010	10	10

Notes: Data collected by the authors. SB means standards-based, EOC means end of course, DOS means depends on subject, RE means Regents examination, and N/A means not available.

B Additional Robustness Checks

Table B.1: Effect of HSEE on Graduation Rate (All States), Non-Parametric Specifications

Variables	Graduation Rate			
	Coefficient	Standard Error	Coefficient	Standard Error
	[1]	[2]	[3]	[4]
$(t - t_i^*)$	0.121*	(0.045)		
$(t_i^* - 10)$			-1.710	(1.126)
$(t_i^* - 9)$			-2.020	(1.143)
$(t_i^* - 8)$			-2.045	(1.161)
$(t_i^* - 7)$			-1.793	(1.226)
$(t_i^* - 6)$			-1.433	(1.231)
$(t_i^* - 5)$			-1.638	(1.163)
$(t_i^* - 4)$			-1.834	(1.268)
$(t_i^* - 3)$			-2.551*	(1.237)
$(t_i^* - 2)$			-3.617**	(1.315)
$(t_i^* - 1)$			-3.644**	(1.069)
(t_i^*)	-2.227**	(0.657)	*Base Category*	
$(t_i^* + 1)$	-2.487**	(0.765)	-3.379**	(1.023)
$(t_i^* + 2)$	-2.370**	(0.912)	-3.193**	(1.082)
$(t_i^* + 3)$	-2.324*	(1.073)	-3.076*	(1.197)
$(t_i^* + 4)$	-1.857	(1.197)	-2.487	(1.427)
$(t_i^* + 5)$	-1.433	(1.148)	-1.985	(1.342)
$(t_i^* + 6)$	-0.754	(1.341)	-1.185	(1.520)
$(t_i^* + 7)$	-1.015	(1.460)	-1.325	(1.638)
$(t_i^* + 8)$	-1.611	(1.543)	-1.828	(1.739)
$(t_i^* + 9)$	-1.164	(1.759)	-1.285	(2.040)
$(t_i^* + 10)$	-1.384	(1.543)	-1.405	(1.784)
$(t_i^* + 11)$	-1.200	(1.648)	-1.101	(1.871)
$(t_i^* + 12)$	-0.026	(1.844)	0.195	(2.024)
$(t_i^* + 13)$	0.276	(1.644)	0.547	(1.793)
$(t_i^* + 14)$	1.383	(1.123)	1.747	(1.210)
$(t_i^* + 15)$	2.143	(1.237)	2.606*	(1.282)
Intercept	73.554***	(0.267)	74.734***	(0.521)
State FE	YES		YES	
Adjusted R^2	0.790		0.785	
N	1,200		1,200	

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Standard errors are clustered at the State level and robust to heteroskedasticity and serial correlation.

Table B.2: EFFECT OF HSEE ON ACHIEVEMENT (ALL STATES), NON-PARAMETRIC SPECIFICATIONS

Variables	NAEP Scores (8^{th} Grade Math)			
	Coefficient	Standard Error	Coefficient	Standard Error
	[1]	[2]	[3]	[4]
$(t - t_i^*)$	0.841***	(0.055)		
$(t_i^* - 10)$			-9.300*	(3.984)
$(t_i^* - 9)$			-7.257*	(3.604)
$(t_i^* - 8)$			-8.502*	(3.678)
$(t_i^* - 7)$			-5.884	(4.476)
$(t_i^* - 6)$			-5.366	(3.908)
$(t_i^* - 5)$			-5.266	(3.658)
$(t_i^* - 4)$			-0.776	(3.103)
$(t_i^* - 3)$			-0.650	(4.989)
$(t_i^* - 2)$			-0.874	(3.729)
$(t_i^* - 1)$			-2.563	(4.482)
(t_i^*)	-0.625	(2.057)	*Base Category*	
$(t_i^* + 1)$	0.735	(1.336)	-0.198	(5.730)
$(t_i^* + 2)$	-0.538	(1.194)	0.321	(3.915)
$(t_i^* + 3)$	3.577**	(1.032)	5.981	(3.512)
$(t_i^* + 4)$	0.348	(1.194)	3.465	(3.403)
$(t_i^* + 5)$	3.325*	(1.097)	5.169	(4.195)
$(t_i^* + 6)$	1.748	(1.334)	5.577	(4.065)
$(t_i^* + 7)$	3.018**	(1.045)	8.463*	(3.381)
$(t_i^* + 8)$	1.233	(1.297)	7.050	(3.838)
$(t_i^* + 9)$	2.565	(1.838)	8.611*	(4.150)
$(t_i^* + 10)$	1.582	(1.304)	7.836	(4.069)
$(t_i^* + 11)$	2.696	(1.814)	10.809**	(3.585)
$(t_i^* + 12)$	0.525	(1.254)	8.402	(4.185)
$(t_i^* + 13)$	1.752	(1.702)	10.004*	(3.822)
$(t_i^* + 14)$	-0.016	(1.097)	7.412	(4.135)
$(t_i^* + 15)$	2.294	(1.457)	11.218**	(3.912)
Intercept	270.055***	(0.368)	275.612***	(1.673)
State FE	YES		YES	
Adjusted R^2	0.914		0.617	
N	457		457	

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Standard errors are clustered at the State level and robust to heteroskedasticity and serial correlation.